DOCUMENT RESUME

ED 459 194 TM 033 480

AUTHOR	Stecher, Brian; Chun, Tammi
TITLE	School and Classroom Practices during Two Years of Education Reform in Washington State. CSE Technical Report.
INSTITUTION	Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.; California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	CSE-TR-550
PUB DATE	2001-11-00
NOTE	43p.
CONTRACT	R305B60002
AVAILABLE FROM	National Center for Research on Evaluation, CRESST/CSE, Graduate School of Education & Information Studies, University of California, Los Angeles, CA 90095-1522. Tel: 310-206-1532.
PUB TYPE	Reports - Research (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Achievement Gains; Educational Assessment; *Educational Change; Elementary Secondary Education; *Principals; *State Standards; Student Evaluation; Teacher Surveys; *Teachers
IDENTIFIERS	*Reform Efforts; *Washington

ABSTRACT

In 1999 and 2000 RAND researchers surveyed Washington state principals and teachers to understand how the state's education reforms, which included standards and assessments, were being enacted at the school level. In 1999, researchers selected a stratified random sample of about 70 elementary schools and 70 middle schools, and in spring 2000 they surveyed about 140 principals and about 200 fourth-grade and seventh-grade teachers. Teachers and principals reported that they understood the reform and were making changes to support it. However, the changes were not uniform and they were occurring gradually. The teacher surveys, which focused on curriculum and instruction in mathematics and writing, revealed changes in the ways these subjects were taught that suggested the state tests were more salient of the teachers than the state standards. Furthermore, there was a mixed pattern of relationships between reported teaching practices and school-level test scores. An appendix contains nine supplemental tables. (Contains 21 tables and 22 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

CRESST

National Center for Research on Evaluation, Standards, and Student Testing

ED 459 194

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

K. Hurst

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

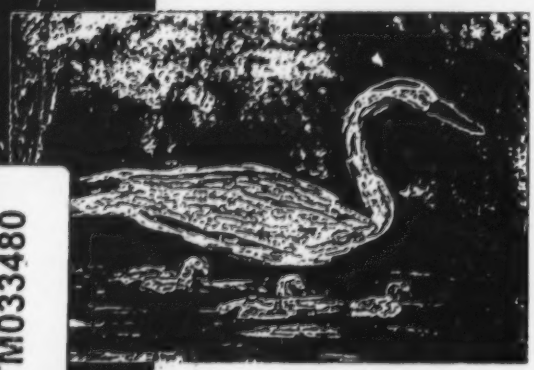
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

School and Classroom Practices During Two Years of Education Reform in Washington State

CSE Technical Report 550

Brian Stecher and Tammi Chun

National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/ RAND



TM033480

UCLA Center for the Study of Evaluation

In Collaboration With

BEST COPY AVAILABLE

School and Classroom Practices During Two Years of Education Reform in Washington State

CSE Technical Report 550

Brian Stecher and Tammi Chun

National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/ RAND

November 2001

Center for the Study of Evaluation
National Center for Research on Evaluation, Standards, and Student Testing
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095-1522
(310) 206-1532

Project 1.5 The Effects of Standards-Based Assessments on Schools and Classrooms
Brian Stecher, Project Director, CRESST/RAND

Copyright © 2001 The Regents of the University of California

The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number R305B60002, as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

The findings and opinions expressed in this report do not reflect the positions or policies of the National Institute on Student Achievement, Curriculum, and Assessment, the Office of Educational Research and Improvement, or the U.S. Department of Education.

SCHOOL AND CLASSROOM PRACTICES IN WASHINGTON STATE DURING TWO YEARS OF EDUCATION REFORM

Brian Stecher and Tammi Chun
National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/ RAND

Abstract

In 1999 and 2000 RAND researchers surveyed Washington state principals and teachers to understand how the state's education reforms, which included standards and assessments, were being enacted at the school level. Teachers and principals reported that they understood the reform and were making changes to support it. However, the changes were not uniform, and they were occurring gradually. The teacher surveys, which focused on curriculum and instruction in mathematics and writing, revealed changes in the way these subjects were taught that suggested the state tests were more salient to the teachers than the state standards. Furthermore, there was a mixed pattern of relationships between reported teaching practices and school-level test scores.

During the past five years, most states and many districts have undertaken standards-based educational reforms. One way in which these efforts differ from earlier reforms is that they involve the adoption of content and student performance standards—explicit benchmarks of what students should know and be able to do. Another way the reforms differ from the past is that they seek to promote "higher" standards that emphasize, among other things, critical thinking and problem-solving skills. In addition, many of these reforms also include accountability systems with tests linked to the standards, and rewards and sanctions linked to school or student performance on the tests.

By the year 1999, 49 of 50 states had established student content or performance standards (Jerald, 2000). In addition, 48 states are developing or have implemented standards-based assessments, and 33 states have accountability measures that set performance goals for schools and school districts holding students, teachers, school administrators and/or district administrators responsible for student performance on the tests (Goertz, Duffy, & Le Floch, 2001). A primary assumption of these reforms is that articulating standards and establishing incentives to meet the standards will motivate improved teaching and learning. Furthermore, many of

these reforms use tests to measure student performance toward the standards and use test scores to judge school success.

Washington state is in the midst of a decade-long standards-based reform of its educational system. This paper presents the findings of research on the impact of the Washington education reform on school and classroom practices, as reported by principals and teachers. We begin with a brief description of standards-based reform in Washington. That is followed by a discussion of our survey research methods. Our findings highlight the transitional nature of the Washington reform, evidenced by variations in educators' knowledge of the reform and changes in school and classroom practices. Finally, we discuss our overall conclusions and the strengths and weaknesses of our approach.

Background

Since 1995, a team of researchers from RAND and the University of Colorado at Boulder has been studying the impact of standards-based reform on schools and classrooms. Earlier, the team conducted surveys and case studies examining the effects of assessment-driven reform in Kentucky (Barron, Borko, McIver, Stecher, & Wolf, 1998; Barron & Stecher, 1999; Borko & Elliott, 1998, 1999; Borko, Elliot, & Uchiyama, 1999; Koretz, Barron, Mitchell, & Stecher, 1996; McIver & Wolf, 1999; Stecher & Barron, 1999; Stecher, Barron, Kaganoff, & Goodwin, 1998; Wolf & McIver, 1999). In 1998-99, the team turned its attention to Washington's reforms to contrast findings in two states with distinguishing features in their assessment and accountability systems.

Washington provides appropriate conditions for studying the effects of reform. The state policies established to support the attainment of the state's learning goals mirror the current educational reforms nationwide. With the passage of the Student Learning and Improvement Act of 1993 (also known as the "Education Reform Act"), Washington state committed itself to an education system that promotes rigorous standards for its students. The goals of the reform are that all students shall

1. read with comprehension, write with skill, and communicate effectively and responsibly in a variety of ways and settings;

2. know and apply the core concepts and principles of mathematics; social, physical and life sciences; civics and history; geography; arts; and health and fitness;

3. think analytically, logically, and creatively, and integrate experience and knowledge to form reasoned judgments and solve problems;

4. understand the importance of work and how performance, effort, and decisions affect future career and educational opportunities. (Student Learning and Improvement Act, 1993)

Washington's reform employs many of the same policy levers used by other states: a set of content-based expectations for learning, tests that are intended to measure performance on standards, professional development to build capacity to teach to the standards, and incentives. Washington's standards are known as the Essential Academic Learning Requirements—EALRs, and the state's testing program is know as the Washington Assessment of Student Learning (WASL). Washington also devoted resources to training teachers to develop more effective classroom-based assessments as a counterpart to the statewide, standardized WASL. They hoped teachers would develop their own high-quality, classroom-based assessments to provide immediate feedback about student performance for instructional improvement. One feature that distinguishes the WASL from previous state tests is that it attempts to measure more critical thinking and problem-solving skills by requiring open-ended responses and essays in response to writing tasks (prompts).

Washington's reform has also been more gradual than reforms in other states, such as California, Kentucky, Maryland and Texas. Rather than implementing the entire system at once, Washington is rolling out the components in steps, having begun with standards and capacity-building in 1995, followed by tests in core subjects of mathematics, reading, writing and communication in 1997. School and student-level accountability and tests in other subjects (social studies, science, health and fitness, fine arts) are scheduled for full implementation by 2008.

Like students in many other states with challenging standards, Washington's students have performed disappointingly on the state tests, as shown in Figure 1. Washington began administering the WASL on a voluntary basis in 1997. The tests were given in Grade 4 in reading, writing, mathematics and communication. In the following years, the tests became mandatory for fourth grade and were introduced on a voluntary basis in seventh grade. For both grade levels, initial results from the WASL in mathematics, reading, and writing showed that fewer than 50% of students met the standards. In fact, in mathematics, only 20% of students met the standards in the first year of test administration. Over time, scores have increased in

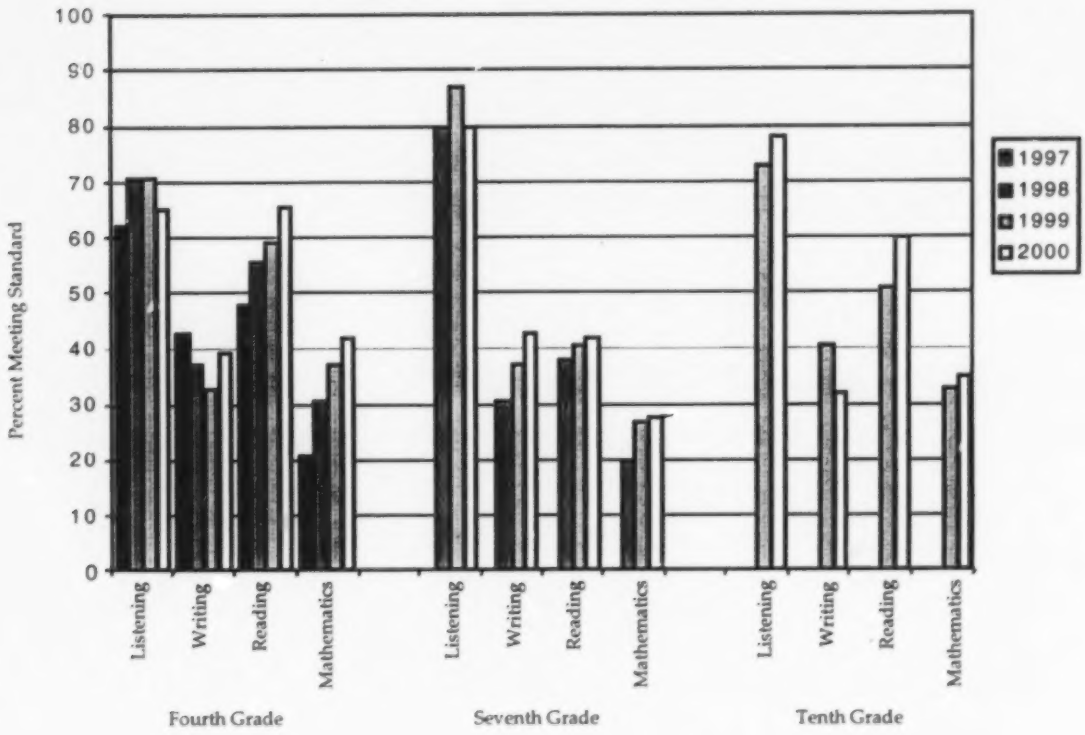


Figure 1. Washington Assessment of Student Learning (WASL) scores for students in Grades 4, 7, and 10.

some subjects and decreased in others. It is common to see score increases in the early years of a new test as teachers and students become familiar with the test format and content emphases. When high stakes are attached to test performance, it is also common to find teachers engaged in test preparation activities designed to improve scores.

Methods

In 1999 and again in 2000, RAND researchers surveyed Washington principals and teachers to understand how the state's education reforms are being enacted at the school level. The surveys included questions about educators' knowledge of the reforms, local policy changes made in response to reforms, changes in classroom practice, activities that prepare students for the tests, and educators' opinions about the reforms. The survey responses provide two consecutive snapshots of opinion and practice in Washington's schools in 1999 and 2000. The results are a valuable resource for understanding the extent to which the standards-based reforms are

being enacted. Herman, Klein, and Abedi (2000) describe such data as "proximate indicators of progress and an important goal in itself" (p. 17).

Sampling

Survey procedures for 1999 are described in detail in Stecher, Barron, Chun, and Ross (2000), and the same procedures were followed in 2000. In brief, each year we selected a stratified random sample of 70 elementary schools and 70 middle schools. We excluded from the sample schools with fewer than 20 students in the tested grade and schools with recent changes in their service areas. There are approximately 1,000 schools in Washington that include Grade 4, and approximately 500 that include Grade 7. Thus, our samples contain about 7% and 14% of the population of schools, respectively. The sample sizes were dictated by resource constraints, but they are not unreasonable for survey research. In the spring we surveyed about 140 elementary and middle school principals and about 400 fourth- and seventh-grade teachers from these schools (see Table 1). All principals were included in the study. In small schools, all teachers in the target grades and subjects (fourth grade, seventh-grade mathematics and seventh-grade writing) were included in the study. In large elementary schools with more than three teachers, a random sample of three teachers was selected. In large middle schools with more than two math or writing teachers, random samples of up to two writing teachers and up to two math teachers were selected.

School cooperation was solicited by telephone in January and February, and names of teachers were obtained during these calls. In March and April principals and sampled teachers were sent surveys by mail. Enclosed with the survey was a letter from the Superintendent of Public Instruction urging respondents to cooperate, a copy of the survey to be completed, a return envelope, and a ten-dollar gift certificate for purchasing books or other instructional materials. A second

Table 1

Survey Samples and Response Rates

	1999		2000	
Respondent	Sample size	Response rate	Sample size	Response rate
Elementary principals	70	75.7%	67	80.5%
Grade 4 teachers	179	74.9%	185	75.7%
Middle school principals	70	78.6%	64	76.6%
Grade 7 writing and mathematics teachers	221	64.7%	208	72.1%

mailing was sent to all nonrespondents about six weeks later. It contained a cover letter and the survey but not the gift certificate. As Table 1 shows, the response rates for both groups of respondents were high in both years. In 2000, 108 principals (82% response rate) and 299 teachers (76% response rate) returned completed surveys.

The median tenure for principals who responded in 2000 was 5 years for elementary and 3 years for middle school principals. The median tenure for fourth-grade teachers and seventh-grade writing teachers was 13 years and for seventh-grade mathematics teachers, 12 years. Overall, teachers had acquired one half of their teaching experience at their current school. About one half of the teachers had master's degrees. Teachers in the sample resembled the teachers in the state as a whole in terms of experience and education level, based on recent research conducted by the legislature (State of Washington Joint Legislative Audit and Review Committee, 1999, pp. 34-35).

In terms of subject matter, almost all fourth-grade teachers in the sample were responsible for teaching all subjects, including reading, writing, mathematics, communication, social studies and science. About one third did not teach art, and about one third did not teach health. Almost all teachers identified as seventh-grade writing teachers also taught reading and communication, and two thirds also taught social studies. A small percentage of the seventh-grade mathematics teachers (10-20%) also taught other subjects in addition to mathematics.

Surveys

The Washington teacher surveys were similar to surveys the project developed and administered in Kentucky in 1997-98; however, they were modified to reflect the language and priorities of the Washington reform. The teacher surveys collected information about teachers' familiarity with and opinions about the state reforms, their participation in professional development, and their classroom practices. In terms of practice, the surveys asked teachers about their allocation of time to different subjects, their teaching strategies, and the topics they covered within writing and mathematics. Additionally, teachers reported recent changes in instruction and the major factors that influenced instructional changes, including the state reforms. Questions about the content of the writing and mathematics curriculum were based on the Washington Essential Academic Learning Requirements (EALRs).

The principal surveys focused on education reform at both the district and school levels. They included questions about curriculum, standards, assessments, and accountability at the district level. The surveys also included questions about implementation (e.g., how the respondent learned about the reform, whether he/she endorsed its principles, etc.), impact (e.g., changes made as a result of the reform, factors that were most influential, etc.), and testing (e.g., test preparation practices) at the school level.

For 2000, we modified the surveys to reflect changes in Washington's education policies (e.g., accountability system, Washington Reading Corps). We deleted some items from the 1999 survey because they did not discriminate well, and we added new items that focused on additional questions of interest. Most of the items relating to teacher background and to mathematics and writing curriculum and instruction were the same in both years.

Data Analysis

For most questions on the principal and teacher surveys, we computed frequency distributions of responses at each point on the response scale. We often combined the top or bottom two categories for reporting, but not if this obscured interesting differences. For questions requiring a numeric response, means and standard deviations were calculated. Because we sampled teachers in the larger schools, we weighted teachers' responses to obtain results that reflected all teachers in Washington (fourth-grade teachers, seventh-grade writing teachers, and seventh-grade mathematics teachers). The weight assigned to each teacher was the product of the inverses of the probability that the school would be selected, the probability the teacher would be selected, and probability that the sampled individuals would participate (complete the survey).

It was not our purpose to draw direct comparisons between groups of teachers or principals, and we do not focus much attention on testing the significance of differences between specific groups of principals or teachers. Instead we focus on differences that seem large enough to be of practical importance. As a general rule of thumb, a difference between two percentage estimates of 15-20 points would be large enough for statistical significance at the 0.05 level (without a correction for multiple comparisons), and we used this significance level informally as a guideline for interpreting results.

To identify a more parsimonious and robust set of variables for describing classroom practices, we conducted principal components factor analyses of the teaching practice items in mathematics and writing, including all the measures of content coverage, teaching strategies and student activities. We conducted these analyses separately for fourth- and seventh-grade teachers, and the resulting factor structure was very similar for both grades. We selected eleven mathematics factors and six writing factors that were similar across grade levels, and we defined new composite variables as the means of the items loading most strongly on each factor. Table 2 lists the composite mathematics and writing variables. The list of items comprising each variable, and the reliabilities of each scale are listed in Tables A.1 and A.2 in the Appendix. We used a similar procedure to derive composite background and opinion variables from the teacher and principal surveys. The items that comprise each composite variable are listed in Tables A.3 and A.4 of the Appendix.

To investigate the similarity of practices across schools, we partitioned the schools in our sample into two groups based on student background factors. The percent of minority students and the percent of students eligible for free or reduced-price lunch were used as the classification variables, and the state median values for elementary schools and for middle schools were used as the cut points. We computed the average score on each derived variable for each school and then

Table 2

Mathematics and Writing Composite Variables

Mathematics	Writing
Mathematical problem solving	Writing process
Relating mathematics to other fields	Conference and assess writing
Constructive mathematics practices	Focus on topic, audience and purpose
Focus on mathematics EALRs and rubrics	Rubric-based writing
Focus on measurement	Practice WASL in writing
Practice WASL in mathematics	Focus on writing conventions
Demonstrating mathematics	
Peer interaction in mathematics	
Writing about mathematics	
Speed drills	
Focus on number sense	

Note. The last three mathematics variables are defined by single items. EALRs = Essential Academic Learning Requirements. WASL = Washington Assessment of Student Learning.

compared the means of these values for high- and low-minority schools and high- and low-income schools.

Finally, we used multiple regression analysis to investigate the relationship between WASL scores and school practices and principal and teacher opinions as reported on the principal and teacher surveys. We obtained school-level WASL scores and student demographic information for 1998-99 and 1999-2000 from the Office of the Superintendent of Public Instruction (OSPI). The data file included the number of students tested in each subject, scaled WASL scores, and the number achieving the standard in each subject. The file also included student demographic information at the school level, including race/ethnicity and eligibility for free or reduced-price lunches. The question we examined was whether school practices and opinions of principals and teachers were significantly related to student achievement, controlling for differences in school size and student demographics.

Our experience in 1999 suggested that background factors were the strongest predictors of WASL scores and that few other measured variables were significant. Lacking a strong basis for identifying variables to include in these models, we included all the derived variables in the model and used stepwise regression analyses to identify significant predictors. We focused on the subjects covered in the surveys—writing and mathematics,—and we estimated separate models at each grade level. Included in the analyses were the derived principal variables, the derived teacher opinion variables, and the derived teacher practice variables for the subjects under investigation.

Results

Surveys in 1999 and 2000 revealed that educators are learning about the Washington reform and making changes to implement it. Changes are widespread—in educators' understandings, district and school policies, and teachers' practice. However, the changes are not uniform and happen gradually. We did not find educators adopting the reform overnight. The surveys revealed a reform in transition—that is, one in which educators had acquired knowledge about the reform and were gradually making changes in practice to support it.

Learning About the Reform

Washington educators continue to learn about the reform through inservice training and professional development. The most common activity undertaken by

almost all schools to learn about the reform was to hold staff meetings that focused on WASL issues. Most schools also focused their inservice training and professional development on WASL-related activities. For both teachers and principals, approximately two thirds or more of teachers reported that their professional development emphasized the EALRs and performance benchmarks, WASL, and aligning curriculum with the EALRs. One half to two thirds of teachers reported that their professional development also emphasized reading, writing, mathematics and communication/listening content. Educators also reported that they understood the reform components. In particular, virtually all principals and most teachers indicated that they understood the EALRs, WASL, and alignment of curriculum and instruction well, though more principals than teachers (by about 10 percentage points) reported understanding the EALRs well. Less well understood were the newer accountability provisions of the reform (e.g., the second-grade reading accuracy and fluency assessment, the fourth-grade reading improvement goals, and the Academic Achievement and Accountability Commission).

It is worth noting that classroom-based assessments were less prominent teachers' and principals' views of the reform than the high-profile standardized test (WASL) or the standards (EALRs). Although two thirds of the teachers believed that better classroom-based assessments would lead to better performance on the WASL, fewer teachers understood classroom-based assessments than understood the WASL or EALRs. About one half of the teachers reported that they understood classroom-based assessments well or very well compared with more than 80% who reported that they understood the WASL and EALRs. This is consistent with the fact that teachers had fewer opportunities to learn about classroom-based assessment than about the other aspects of the reform. Although one half of the principals reported that their schools' professional development emphasized classroom-based assessment, this figure was lower than the percentage that emphasized the WASL, EALRs or learning about specific content areas. Less than one third of the teachers reported that their own professional development emphasized classroom-based assessment, whereas more than two thirds of teachers indicated their professional development emphasized WASL and EALRs. About 75% of teachers felt pressure for students to perform well on classroom-based assessments, but equal numbers felt pressure regarding students' classroom grades, and almost all teachers felt pressure for students to perform well on the WASL.

Principals and teachers were generally supportive of the reform, though more principals than teachers supported the reform and expected it to endure. Three fourths of principals thought that the goals of the reform are attainable; about 70% of principals thought that the time spent administering the state tests is well spent and that these reforms will persist beyond five years. Teachers were less positive. Sixty percent of seventh-grade teachers, but only 40% of fourth-grade teachers agreed that the goals of the reform are attainable. Additionally, 20% of fourth- and seventh-grade teachers strongly disagreed that the goals are attainable. Only 40% of the teachers thought that the information learned from state testing is worth the time spent administering the test, and 70% of teachers thought that this reform (WASL, EALRs) will be replaced in four or five years.

School and District Policy Changes

Schools and districts are changing policies to support the Washington education reform. In 2000, 90% of schools had developed school plans to improve performance on WASL. This represents an increase of 18 percentage points since 1999 in the number of middle schools reporting school planning as a strategy for improving WASL (Stecher et al., 2000). About one half of the schools added summer school to increase instructional time, and many said they did this in response to WASL. To improve WASL scores, about one half of the schools also changed their school schedules to add time for reading.

Nationally, state testing programs have reduced the role of local tests in assessment and accountability systems. Yet, districts often supplement state tests with local testing programs (Goertz et al., 2001). This appears to be true in Washington, as well. In 2000, about 40% of elementary school principals and 60% of middle school principals reported that their districts were changing grades, subjects or content tested to align with EALRs. The percentages were slightly higher in 1999. About one half of principals reported that districts continued to introduce new local assessments in 2000; compared to 1999, 10 to 25 percentage points fewer principals reported new local assessments. Only about 20% of principals reported that their districts were phasing out district assessments. Thus, it appears that the testing burden on Washington schools and students continues to increase.

Classroom Practice Changes

The surveys revealed changes in teachers' allocation of time as well as in their curriculum, pedagogy and student learning activities. The changes in use of instructional time are dramatic, but the changes in teacher practices are more subtle.

There has been a shift in instructional time from non-tested to tested subjects. Fourth-grade teachers who teach all subjects reported that they are spending two thirds of their instructional time (median of 16 hours per week) on the four tested subjects. Teachers spent most time on reading, followed by mathematics, and then writing. The median teacher spent 5 hours per week on reading, 5 hours per week on mathematics, and 4 hours per week on writing. Even here, there was considerable variation among teachers, with two thirds of the teachers falling in the range of 2 to 8 hours per week on reading, 3 to 7 hours per week on mathematics, and 2 to 6 hours per week on writing. In the other subjects, the median teacher spent only 1 to 2 hours per week per subject. Figure 2 shows changes in teachers' allocation of class time. In both 1999 and 2000, teachers reported increasing the time they spent on reading, writing, mathematics and communication and decreasing the time on untested subjects of social studies, science, arts, and health and fitness. This occurred despite the fact that there are EALRs for all subjects. This finding was consistent over the two years of the survey, though data from the 2000 survey suggested that fewer teachers increased time spent on reading and writing in 2000 compared to 1999, so the reallocation of time may be reaching an equilibrium (Stecher et al., 2000).

Greater attention to the WASL-tested subjects is one way in which the WASL's influence seems to be greater than the influence of the EALRs. Other evidence that supports this conclusion comes from teachers' reports about curriculum alignment and the factors that influence their practice. In 1999, about 90% of teachers said that EALRs and WASL had some influence on changes in their mathematics and writing teaching, but the WASL's influence was stronger. Almost two thirds of fourth-grade teachers and about one half of seventh-grade mathematics teachers reported that the WASL extended-response-type items strongly influenced their mathematics instruction. About 40% fewer said that EALRs strongly influenced their instruction. The differences were similar but less dramatic in writing. Additionally, more teachers and principals reported that local standards and curriculum were well aligned with EALRs in tested subjects than non-tested subjects.

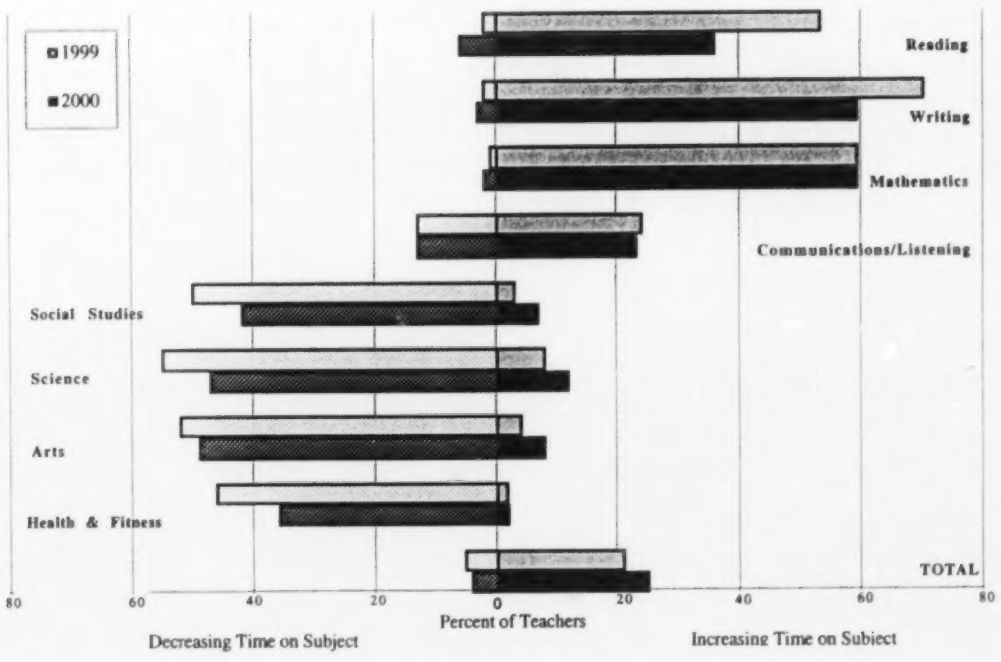


Figure 2. Reallocation of instructional time, fourth-grade teachers, 1999 and 2000.

Writing

Teachers reported gradual changes in their writing instruction over the last three years. Most often they reported little or no change in the most common aspects of their practice, but increases in elements of the curriculum or teaching practices that were previously addressed infrequently. This suggests that instruction is changing on the margins, as teachers add content or strategies to their existing lessons or pedagogy. The greatest changes were related to the distinctive aspects of the WASL in writing, which requires students to respond to two writing prompts in different modes or genres (narrative and expository for fourth grade, and persuasive and expository for seventh grade). The students' WASL writing samples are scored with a rubric that emphasizes genre (topic, audience and purpose) and writing conventions.

The frequency of the derived writing practice variables reflects these two emphases. Table 3 shows the average frequency of each writing composite variable. Attention to writing conventions (Focus on conventions) is the most frequent

Table 3

Mean Frequency Scores on Writing Practice Composite Variables in 2000

Composite variable	Fourth grade	Seventh grade
Focus on conventions	4.1	4.3
Writing process	3.6	3.8
Focus on topic, audience and purpose	3.6	3.6
Conference and assessment	3.2	3.3
Rubric-based writing	3.2	3.4
Practice WASL	3.1	3.1

Note. 1 = zero times per year (never); 2 = 1-5 times per year (1-2 times per semester); 3 = 6-30 times per year (1-2 times per month); 4 = 31-80 times per year (1-2 times per week); 5 = more than 80 times per year (almost daily). WASL = Washington Assessment of Student Learning.

component of writing instruction in both grades, but all the composites have average frequencies above 3.0. (This is not the case in mathematics, as will be shown below.) On average, writing conventions are covered more than one to two times per week. The next most frequent practices focus on the writing process and on genre, emphasizing topic, audience and purpose. Both of these are closely associated with the WASL. Rubric-based writing strategies and specific practice on WASL occur somewhat less frequently, but both occur more than one to two times per month, on average.

It is possible to sense the evolution of practice by looking at teachers' reports of changes in their curriculum emphasis. It appears that the emphasis on writing process and writing conventions preceded WASL, because few teachers reported increasing their focus on these areas. Table 4 shows the percentage of teachers who covered each of the writing EALRs at least weekly and the percentage of teachers reporting increased coverage of the writing EALRs since 1999. The greatest increase in writing coverage related to genre: writing for different purposes, in a variety of forms, for different audiences, and using a style that is appropriate to the audience and purpose. One third to one half of teachers at both grade levels increased the frequency with which they covered these aspects of writing. This change is quite consistent with the elements of writing that are emphasized in the WASL scoring rubrics.

Teachers increased their emphasis on genre, particularly the genre tested by the WASL at their respective grades. The EALRs identify four genres: narrative, persuasive, expository and descriptive. In general, most students were exposed to all

Table 4

Percent of Teachers Covering Each Writing EALR in 2000

Writing EALRs	Daily or weekly coverage		Increase in coverage since 1999	
	Grade 4	Grade 7	Grade 4	Grade 7
Application of writing conventions	84	84	34	40
Writing process: draft	71	70	39	41
Writing process: edit	65	70	38	37
Writing process: pre-write	65	69	39	40
Writing process: revise	63	68	40	44
Development of concept and design	47	36	56	42
Seek and offer feedback to other students	46	38	37	33
Genre: write for different purposes	45	67	45	57
Genre: style appropriate to audience and purpose	41	54	38	42
Genre: write for different audiences	36	45	41	42
Genre: write in a variety of forms	35	64	41	53
Writing process: publish	35	49	29	36
Assessment of students' own strengths and needs for improvement	33	37	39	34
Write for career applications	12	18	16	18

Note. EALRs = Essential Academic Learning Requirements.

four genres during the year, but not with equal frequency. About two thirds of fourth-grade teachers assigned writing in each genre except persuasive writing at least monthly, and about one half of seventh-grade writing teachers assigned each genre except persuasive writing at least monthly.

In 1999, OSPI decided that the Grade 4 WASL writing prompts would require narrative and expository writing, and the Grade 7 WASL writing prompts would require persuasive and expository writing. The attention to writing in particular genres seems to have shifted since then. First, there has been a general increase in attention to writing. Overall, one half of the surveyed teachers increased their use of writing prompts in the classroom. At least one third of fourth-grade teachers reported increasing writing assignments in all genres. Second, there have been greater increases in the tested genres. About one half of fourth-grade teachers increased their assignment of expository writing, and about one third of them have increased their assignment of narrative, persuasive and descriptive writing. In seventh grade, about one half of the teachers increased the frequency of assignments requiring expository and persuasive writing.

Like the writing curriculum, writing instruction has also changed mildly, as shown in Table 5. Reading orally to students and explaining language mechanics are reported as the teaching strategies used at least weekly by three quarters or more of the teachers. However, the areas of greatest change are in the strategies used less frequently. Only about one third to one half of the teachers used rubrics in teaching writing either by commenting on student work in terms of WASL rubrics or teaching rubric-based approaches to writing (e.g., Six-Trait writing); but about one half of the teachers reported increased use of these strategies.

In addition to changes over time and differences between the grades, we also found some differences in writing curriculum and instruction associated with the ethnic composition of the schools' student population (see Table 6). These differences were significant in the seventh grade but not the fourth grade. Seventh-grade students in high-minority schools were exposed to almost all of the writing practice elements more frequently than seventh-grade students in low-minority

Table 5

Percent of Teachers Using Selected Teaching Strategies in Writing in 2000

	Daily or weekly use teaching strategy		Increase in coverage since 1999	
Writing teaching strategies	Grade 4	Grade 7	Grade 4	Grade 7
Read orally to students	94	73	18	21
Explain correct usage of grammar, spelling, punctuation and syntax	81	90	31	33
Give examples of choosing appropriate words to describe objects or experiences	59	64	37	33
Suggest specific revisions to student writing	58	57	37	29
Teach Six-Trait or other rubric-based approach to writing	55	57	46	46
Use examples to discuss the craft of an author's writing	55	57	28	30
Provide a prompt to initiate student writing	54	55	49	48
Assess students' writing skills	49	67	37	34
Demonstrate use of pre-writing	49	38	41	32
Provide time for unstructured ("free") writing	45	52	20	17
Conference with students about their writing	34	25	27	19
Provide time for students to conference with each other about writing	34	30	33	36
Show examples of writing in different content areas	34	28	31	32
Comment on student writing in terms of WASL rubrics	30	41	47	59
Write with students on the same assignment	13	14	21	20

Note. WASL = Washington Assessment of Student Learning.

Table 6

Mean Frequency Scores on Writing Practice Composite Variables by School Percent Minority in 2000

Composite variable	Fourth grade		Seventh grade	
	Low[a]	High[b]	Low[c]	High[d]
Writing process	3.7	3.6	3.6	3.9
Conference and assessment	3.3	3.2	3.1*	3.4*
Focus on theme, audience and purpose	3.4	3.3	3.4*	3.7*
Rubric-based writing	3.2	3.3	3.2*	3.5*
Practice WASL	3.1	3.1	3.0	3.2
Focus on conventions	4.2	4.0	4.1*	4.4*

[a]Less than or equal to 18.9% minority. [b]Greater than 18.9% minority. [c]Less than or equal to 15.4% minority. [d]Greater than 15.4% minority. WASL = Washington Assessment of Student Learning.

*$p < .05$.

schools. Unfortunately, the survey does not provide any clear explanation for these differences. Further examination of these data and other information are warranted.

Overall, changes in writing appear to be consistent with incentives created by the Washington education reform, particularly the emphases of the WASL test in writing. Teachers reported increasing their coverage of aspects of writing that are most relevant to the WASL although they did not report commensurate decreases in other aspects of writing.

Mathematics

As with writing, the most frequently occurring elements of the mathematics curriculum are fairly traditional—teacher-led demonstrations and a focus on number sense—but there have been important changes in mathematics over the last two years to make curriculum and instruction more consistent with the goals of the reform.

Table 7 shows the frequency of the 11 major mathematics composite variables. As noted above, teachers conducted demonstrations and taught lessons that focused on number sense more often than any other aspect of mathematics. However, the next most frequent practices were associated with a more standards-based view of mathematics—having students interact with one another and focusing on problem solving. The least frequent practices involved teaching measurement and focusing on the EALRs, and the gap between the most and the least frequent activities is

Table 7

Mean Frequency Scores on Mathematics Practice Composite Variables in 2000

Composite variable	Fourth grade	Seventh grade
Demonstrating mathematics	4.1	4.2
Focus on number sense	4.1	4.0
Peer interaction	3.8	3.5
Problem solving	3.7	3.5
Write about mathematics	3.6	3.5
Practice WASL	3.5	3.3
Conduct speed drills	3.5	1.9
Relating mathematics to other areas	3.4	3.6
Constructive mathematics	3.1	2.8
Focus on measurement	2.9	2.9
Focus on EALRs and rubrics	2.8	2.6

Note. 1 = zero times per year (never); 2 = 1-5 times per year (1-2 times per semester); 3 = 6-30 times per year (1-2 times per month); 4 = 31-80 times per year (1-2 times per week); 5 = more than 80 times per year (almost daily). EALRs = Essential Academic Learning Requirements. WASL = Washington Assessment of Student Learning.

large. More time was spent focusing on WASL than on the EALRs. With one exception, the results are similar for fourth-grade teachers and seventh-grade mathematics teachers. The exception is the infrequent use of speed drills in seventh grade.

Table 8 shows the relative emphasis teachers placed on the five math content areas identified specifically in the EALRs. More than two thirds of the teachers covered number sense at least weekly, far more than any of the other content areas. However, the content of mathematics has been changing. The area of greatest

Table 8

Percent of Teachers Covering Each Mathematics Content EALR in 2000

	Daily or weekly coverage		Increase in coverage since 1999	
Mathematics content EALRs	Grade 4	Grade 7	Grade 4	Grade 7
1.1 Number sense	79	68	33	24
1.2 Algebraic sense	37	57	40	31
1.3 Measurement	27	23	25	16
1.4 Geometric sense	27	20	34	24
1.5 Probability and statistics	22	14	43	33

Note. EALRs = Essential Academic Learning Requirements.

change was in coverage of probability and statistics, where at least one third of the teachers reported increasing time spent on the content area. Many teachers also increased their coverage of algebraic sense and geometric sense. The area of least change was measurement.

Table 9 provides similar information for the mathematics process components of the EALRs. Perhaps the most significant information in the table is the extent of the increase in coverage of all the mathematical processes. For each element, between one third and one half of the fourth-grade teachers reported an increase in coverage since last year. In seventh grade the change was slightly smaller but equally widespread. Clearly, teachers are devoting more attention to these aspects of mathematics. Because of the way the EALRs are constructed, the more basic procedures, such as computation and the application of formulas and algorithms, were not included in the set of processes.

Mathematics teaching practices are changing along with mathematics content. Table 10 shows that traditional methods are the most common, but more reform-oriented practices are used regularly and are growing in popularity. Almost all teachers reported explaining correct solutions and demonstrating new skills at least weekly, and about three quarters reported regularly asking open- response questions with many right answers. This was also the area of greatest change.

Table 9

Percent of Teachers Covering Each Mathematics Process EALR in 2000

	Daily or weekly coverage		Increase in coverage since 1999	
Mathematics process EALRs	Grade 4	Grade 7	Grade 4	Grade 7
3.1 Analyze information	75	70	41	44
3.3 Draw conclusions and verify results	68	56	39	27
2.3 Construct solutions	68	59	29	32
5.3 Relate concepts to real life	65	82	42	34
4.2 Organize and interpret information	62	48	45	32
2.1 Investigate situations	62	58	35	28
5.1 Relate concepts within math	62	66	49	35
2.2 Formulate questions	56	49	46	40
4.3 Represent and share information	54	32	55	45
4.1 Gather information	51	43	50	31
5.2 Relate concepts to other disciplines	43	43	45	27

Note. EALRs = Essential Academic Learning Requirements.

Table 10

Percent of Teachers Using Selected Teaching Strategies in Mathematics in 2000

Mathematics teaching strategies	Daily or weekly use		Increase in use since 1999	
	Grade 4	Grade 7	Grade 4	Grade 7
Explain correct solutions	90	93	28	17
Demonstrate new skill	84	81	22	15
Ask open-response questions with many right answers	79	68	48	42
Explain new concept	75	78	21	11
Assess students' mathematics skills	71	79	24	18
Give examples of real-life applications	70	76	32	25
Conduct speed drills	52	11	16	7
Demonstrate mathematical ideas using constructions, manipulatives, etc.	51	38	31	31

Almost one half of the teachers increased their use of open-response questions, and about one third increased the frequency with which they gave examples of real-life situations in mathematics. Both strategies are consistent with the reform. Fourth-grade teachers were much more likely than seventh-grade teachers to conduct speed drills regularly; one third of seventh-grade teachers actually decreased their use of speed drills. Fourth-grade teachers were also more likely to have students demonstrate ideas using manipulatives than were seventh-grade teachers.

We compared the mathematics curriculum and teaching composite scores between schools serving high and low percentages of minority students, and we found that reform-oriented activities were more common in high-minority middle schools. As shown in Table 11, students in high-minority middle schools were asked to write about mathematics, to relate mathematics to other subjects, and to do hands-on and constructive mathematical activities more frequently than were students in low-minority schools. Students in high-minority middle schools also worked with other students and focused on the EALRs and rubrics in mathematics somewhat more frequently than students in low-minority schools, although the differences were not statistically significant. There were no significant differences at the elementary level. This pattern was not evident when we compared schools on the basis of income level (see Table A.5 in Appendix A).

Table 11

Mean Frequency Scores on Mathematics Practice Composite Variables by School Percent Minority in 2000

	Fourth grade		Seventh grade	
Composite variable	Low[a]	High[b]	Low[c]	High[d]
Demonstrating mathematics	4.0	4.1	4.2	4.2
Focus on number sense	4.0	4.2	3.9	4.1
Peer interaction	3.9	3.8	3.3	3.6
Problem solving	3.8	3.6	3.4	3.6
Write about mathematics	3.8	3.5	3.2*	3.7*
Relating mathematics to other areas	3.4	3.4	3.3*	3.7*
Practice WASL	3.4	3.5	3.2	3.3
Conduct speed drills	3.4	3.5	1.8	2.0
Constructive mathematics	3.1	3.0	2.5**	3.0**
Focus on EALRs and rubrics	2.9	2.7	2.3	2.7
Focus on measurement	2.9	2.9	2.9	2.9

[a]Less than or equal to 18.9% minority. [b]Greater than 18.9% minority. [c]Less than or equal to 15.4% minority. [d]Greater than 15.4% minority. EALRs = Essential Academic Learning Requirements. WASL = Washington Assessment of Student Learning.
*$p < .05$. **$p < .01$.

Preparation for WASL

WASL is prominent in the minds of educators. All principals and teachers feel a moderate to great deal of pressure for their students to perform well on WASL. We asked teachers about the activities used in the classroom to help students do well on WASL tests. Many test preparation activities were evident, and teachers reported spending a fair amount of time, particularly as the testing date approached, preparing for the test.

Some test-related activities have been incorporated into regular instructional practice in writing and mathematics. For example, more than one half of the writing teachers used rubric-based approaches to teaching writing and used open-ended questions in the classroom at least weekly. In mathematics, more than one half of the fourth-grade teachers and more than 40% of the seventh-grade mathematics teachers used open-ended questions in class work and had students practice using WASL-like items at least weekly.

Furthermore, teachers increased the amount of time spent on explicit WASL preparation as the test approached in the spring. Near the beginning of the year, in

November, about one half of the teachers spent 1 to 2 hours a week preparing for the WASL. About one quarter of the teachers did not do any explicit test preparation in November, and fewer than 10% spent an hour a day or more on test preparation. However, as the test approached, teachers increased the amount of time spent preparing for it. In April, with the test imminent, one third of fourth-grade teachers and one fifth of seventh-grade teachers reported spending more than 4 hours per week preparing for the test; fewer than 10% of the teachers reported spending no time on test preparation. Figures 3 and 4 show WASL preparation time reported by mathematics teachers. The results were similar for writing teachers.

The surveys also indicated that teachers pay more attention to the WASL than to the EALRs that the tests are supposed to reflect. We presented teachers with two contrasting viewpoints on addressing the EALRs and WASL and asked them to identify their own approach relative to these two. The first point of view focused on the standards: "I teach the EALRs, and I don't bother with WASL preparation at all. If students master the EALRs, they will do well on the WASL." The contrasting viewpoint focused on the test: "I teach to the WASL, and I make sure my students practice the kinds of questions they will encounter when they take the test. It is

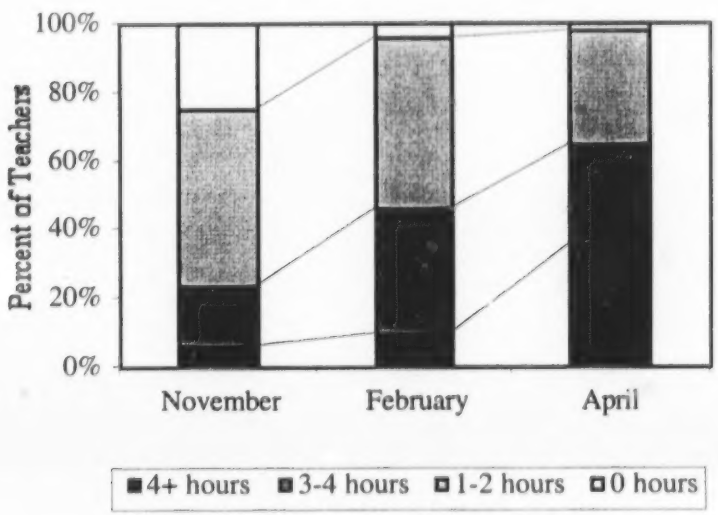


Figure 3. Hours per week spent in fourth-grade classrooms preparing for the WASL test in mathematics in 2000.

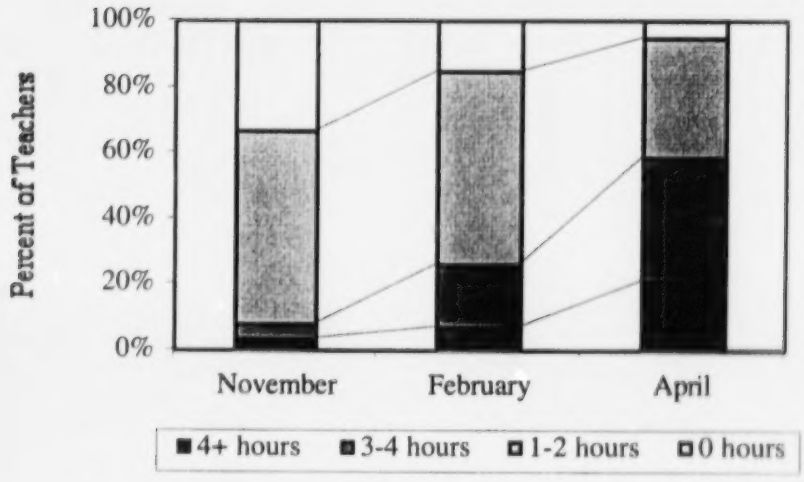


Figure 4. Hours per week spent in seventh-grade mathematics classrooms preparing for the WASL test in 2000.

important for students to master the material on the WASL." Two thirds of teachers identified their teaching as more like "teach[ing] to the WASL," than "teach[ing] the EALRs." Principals responded similarly when asked about the approach they encouraged at their school.

The Relationship Between School and Classroom Practices and WASL Scores

We conducted regression analyses to see whether school and classroom practice data collected in our surveys were associated with WASL scores or WASL score gains. We predicted WASL scores as a function of aggregate student demographic variables, principal opinion composite variables, teacher opinion composite variables, and teaching practice composite variables (i.e., mathematics practices were included in the mathematics model, writing practices were included in the writing model, and no practice variables were included in the reading model). For the most part, the only variables that were significant predictors of WASL scores were aggregate student demographic factors. Table 12 summarizes the significant relationships revealed by the six regression models we constructed for mathematics, writing and reading in Grades 4 and 7. The full results are listed in Tables A.7 to A.9 in the Appendix. In a few cases, attention to specific school or classroom practices was associated with higher WASL scores. For instance, having a curriculum that was aligned with EALRs and devoting more time to relating mathematics to life and to

Table 12

Summary of Regression Results

Variables positively related to 2000 WASL scores	Variables negatively related to 2000 WASL scores
Mathematics	
Percent Asian (4, 7)	Percent free/reduced-price lunch (4, 7)
School curriculum aligned with EALRs (4)	Percent Hispanic (4)
Relating mathematics to life and other fields (4)	School still needs to work on reform (7)
Writing	
Percent Asian (4)	Percent free/reduced-price lunch (4, 7)
Rubric-based writing (4)	Principal support for reform (4)
Focus on writing conventions (7)	School actions in support of reform (7)
Reading	
	Percent free/reduced-price lunch (4, 7)
	Percent Hispanic (4, 7)
	Percent American Indian (4)

Note. Numbers in parentheses are grade levels where factor was significantly related to WASL scores. WASL = Washington Assessment of Student Learning.

other fields were associated with higher math scores in either fourth or seventh grade. Still needing to work on some aspects of the reform (in the opinion of the principal) was associated with lower scores. Similarly, teaching more rubric-based writing was associated with higher WASL writing scores in one grade. These results are encouraging. However, other results are more difficult to understand. For example, it is harder to see why taking more actions to support the reform would be associated with lower writing scores at the seventh-grade level. Overall, we did not find strong evidence that average practices measured by our surveys were directly related to school success on the WASL. We also examined the relationship between school practices and WASL gains from 1999 to 2000 with similar, inconclusive results.

Discussion

Two years of surveys in Washington state reveal that school policies as well as classroom curriculum and instruction are changing in ways consistent with the Washington education reform. However, the changes are gradual, and neither curriculum nor teaching has being transformed overnight. Furthermore, the changes are uneven; some schools are making more progress than others. Some of the

variation between schools is associated with student ethnicity and family income. We also detected a stronger influence of the WASL relative to the EALRs—a difference identified by educators and apparent in their behaviors—that raises some concern about the incentives created by the system and the ultimate direction of the reform. Washington's approach to reform—a slow, decade-long implementation of standards, tests and accountability—has features to recommend to other states.

Gradual Changes

The survey results reinforce the common belief that change takes time. After two years, we find evidence that many teachers are changing their practice, but these changes are gradual. For example, although many teachers reported increases in the amount of class time they devote to mathematics, the median number of hours devoted to mathematics per week did not change from 1999 to 2000. This relationship was true for reading and writing as well—a substantial percentage of teachers increased their coverage, but the overall median amount of time did not change. The greatest changes in these subjects occurred in previously untaught or infrequently taught aspects of the curriculum, such as statistics and probability in mathematics and rubric-based writing. This response by teachers may be appropriate. Attention to the EALRs and/or the WASL may be revealing gaps in instruction—standards that are not taught—and encouraging teachers to include aspects of subjects that were previously untaught.

The evidence also suggests that teachers are adding new instructional strategies in reading, writing, and mathematics without dropping any existing practices. The surveys revealed increases in many aspects of instruction without concomitant decreases in other aspects. This is explained, in part, by increases in the amount of instructional time devoted to these subjects. However, in the long run, this increase is a cause for concern. The untested subjects are scheduled to become tested subjects in a few years. Adding more content and new teaching strategies without making necessary trade-offs is not a sustainable way to respond to the reform.

Uneven Changes

The surveys indicate that school and classroom changes are uneven. While principals and teachers uniformly feel a great deal of pressure for their students to perform well on the WASL, principals' and teachers' actions have been less uniform. The surveys reveal a lot of variation in educators' responses to the education reform. Overall there appears to be progress, but differences between and within schools

remain. These differences, particularly those associated with ethnicity and socioeconomic status of students, need to be studied further. Lasting differences in opportunities to learn will have equity implications.

Improving WASL Scores

One of the goals of this research was to identify changes in practice that were associated with improved outcomes and to use these as a basis for improvement. Overall, teachers' practice appears to be more reform-oriented each year, and, in general, WASL scores are increasing. Yet, this general trend was not reflected in our analyses of specific practice. We did not detect any relationships between curriculum emphases, teaching practice, or educators' beliefs and school-level WASL scores. In 1999 we found that greater curriculum alignment was associated with higher WASL scores (Stecher et al., 2000), but that finding was not replicated this year.

There are several possible explanations for the disappointing finding. First, the survey may have not captured the key aspects of teaching that affect WASL scores. It is possible that the types of curricular and instructional practices related to score changes are not the ones we included in the survey or are not ones that are amenable to measurement using survey techniques. It is also possible that frequency of practice is not as important as quality of practice, something we could not measure with our survey. Another possible explanation for the lack of measurable relationship is aggregation. WASL scores were available only at the school level, so we averaged teachers' responses at each school and compared the average response to the school's WASL-scaled score. Using averaged teacher responses and school-level performance may weaken the association between teachers' actual practice and their students' performance. In addition, having to conduct the analysis at the school level rather than the teacher level effectively reduces our sample size by about two thirds. Reduced sample size decreases the precision of the estimates in the regression model, making it difficult to detect all but the most robust relationships. It is also possible that teachers' reports of their practice do not accurately reflect the practices in their classroom (cf. Cohen, 1990). The fact that we did not find significant effects for most practice variables is not conclusive evidence that they do not exist.

Influence of the Test

As is the case in many other states that have adopted standards-based reforms with high-stakes tests, Washington's state assessment is the most high-profile component of the reform. Teachers and principals feel a great deal of pressure for students to perform well on the WASL. Schools have made institutional and organizational changes, such as instituting summer school, in response to past WASL scores. Teachers have shifted instructional time and focus to subjects tested by the WASL. Most teachers spend time preparing students for the test, particularly as the testing date approaches.

However, this attention to the WASL, rather than the EALRs, which the WASL is supposed to measure, raises concerns about the extent of "teaching to the test" and the generalizability of WASL scores. Tests like WASL cover only a fraction of the domains of performance identified in the standards. The tests are designed to sample student performance, not to exhaustively measure students' mastery of a broad domain. We make inferences from the test content to the domain based on the belief that the test is a random and representative sample of content. Focused preparation on tested format and content that ignores untested content can artificially inflate scores, providing a misleading measure of students' test-taking ability rather than a measure of their attainment of standards.

Research suggests that teachers respond to high-stakes tests by narrowing curriculum and devoting time to test preparation at the expense of broader learning goals (Stecher, 1999; Stecher & Mitchell, 1995). We tried to determine whether this was the case in Washington—that is, whether there was greater attention paid to the content and format of the test than to the broad underlying goals of the reform and the domains of knowledge represented in the standards. The surveys asked teachers a number of questions about the influence of WASL versus EALRs on their practices and on students' test scores. There is no doubt that teachers are responding to the WASL. Many fourth-grade teachers have reallocated most of their teaching time so that 68% of their time is spent on reading, writing and mathematics. This increase in time on tested subjects has been at the expense of other subjects. For example, the role of social studies in the curriculum has been reduced dramatically. Most teachers spend only 1 to 3 hours per week teaching social studies, and one half of the teachers have reduced this allocation. In writing, teachers reported increasing assignments in tested genres and decreasing attention to untested genres, despite the Commission on Student Learning's admonition that the assignment of tested genres to particular

grades was "in no way meant to limit classroom instruction or district and classroom-based assessments" (Ensign, 1999).

The focus on currently tested subjects seems short-sighted because WASL tests in science, social studies, arts, and health and fitness are on the horizon. They are scheduled to be added to the system over the next four to five years. Washington educators believe that increased classroom time is a major factor contributing to increased test scores (Bergeson, Mayo, Fitten, & Bylsma, 2000). Decreasing classroom time in non-tested subjects may lead to lower scores on those tests when they are implemented. Attention to the new tests may lead to reductions in classroom time on the old subjects, causing those scores to drop. Instructional time is a fixed resource, and large shifts due to short-term considerations are likely to have undesirable long-term consequences.

An additional concern is that the new WASL tests will be administered in Grades 5 and 8, instead of Grades 4 and 7. This policy is designed to reduce the testing burden on teachers and students in the fourth and seventh grades. However, the practice also removes the direct incentive for teachers to address standards in all subjects. Previous teacher surveys in Kentucky indicated that teachers' curriculum coverage was directly related to the subjects tested at their grade (Barron & Stecher, 1999). Kentucky students in fourth and seventh grades received more instruction in reading, writing and science, whereas students in fifth and eighth grades received more instruction in mathematics, social studies, and arts/humanities. The consequence of not addressing the existing reallocation of time in fourth grade may be poor student performance on the new fifth-grade WASL tests of science, social studies, arts, and health and fitness due to lack of foundation skills.

Time to Reflect

One noteworthy characteristic of Washington's reform is the gradual implementation of assessments and accountability provisions. The experience of other states suggests that Washington has been wise to implement slowly. The schedule has allowed Washington educators to learn about the reform, and most educators appear to support it. State policymakers have also had the time to address concerns and adapt the reform policies to maintain support for the reform. For example, many educators raised concerns about the difficulty level of the WASL in mathematics for Grade 4. Indeed, only 37% of elementary principals and 19% of fourth-grade teachers agreed that "WASL standards for mathematics are of an

appropriate difficulty for the tested grades." In response, OSPI conducted a study that determined the test was appropriate (Bergeson et al., 2000). Their efforts addressed teachers' concerns. The state has taken its time in other ways. The WASL in science was scheduled for mandatory administration in spring 2001, but implementation was delayed to improve the test. State Superintendent Terry Bergeson said, "I can't put a test out that isn't ready . . . The test needs to be a very good vision of science education [since] it will drive curriculum and teaching" (Abe, 2000).

Other states that have charged ahead have found themselves on the defensive when critics raised concerns about the tests' technical quality or the implications of accountability policies. For example, in 1998 Kentucky legislators replaced the Kentucky Instructional Results and Information System (KIRIS) reforms with the Commonwealth Accountability and Testing System (CATS) in response to concerns and controversy over the reliability of performance tasks on the test, distribution of monetary awards to school personnel for increased test scores, and designation of schools as "low performing" when their scores decreased. Similarly, many states, including Michigan and Florida, are having to back down from tough accountability policies that attach "high stakes"—monetary awards, voucher opportunities, educator employment, diploma endorsements and high school graduation—to test scores because of the political pressure the policies generate. Washington benefits from the opportunity to learn from other states' experiences. Its legislature is currently considering an accountability system and the Certificate of Mastery requirement for graduation—which includes passing WASL in reading, writing and mathematics.

The downside of gradual implementation may be the difficulty of sustaining teachers' enthusiasm for changing over a long period of time. Washington teachers have begun to change their practices, in part due to the current WASL exams. Over time, these practices may become institutionalized. Teachers may resist further changes in practice because they see the results (or lack of results) in current WASL tests, they are exhausted from the changes, or the incentives to make changes in new content are targeted at different grade levels.

References

Abe, D. (2000, October 17). Science part of test on hold. *The News Tribune*. Retrieved October 17, 2000, from http://www.tribnet.com/frame.asp?/new.top_stories/1017b11.html

Barron, S., Borko, H., McIver, M., Stecher, B., & Wolf, S. A. (1998, June). *The effects of standards- based assessment on schools and classrooms.* Session presented at the 1998 CCSSO National Conference on Large-Scale Assessment, Colorado Springs, CO.

Barron, S., & Stecher, B. (1999, April). *Changes in curriculum in response to assessment subject and grade level.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, PQ, Canada.

Bergeson, T., Mayo, C. L., Fitten, R., & Bylsma, P. (2000, September). *Study of the grade 4 mathematics assessment: Final report.* Olympia, WA: Office of the Superintendent of Public Instruction.

Borko, H., & Elliott, R. (1998, November). *Tensions between competing pedagogical and accountability commitments for exemplary teachers of mathematics in Kentucky* (CSE Tech. Rep. No. 495). Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Borko, H., & Elliott, R. (1999). "Hands-on" pedagogy vs. "hands-off" accountability: Tensions between competing commitments for exemplary teachers of mathematics in Kentucky. *Phi Delta Kappan, 80,* 394-400.

Borko, H., Elliott, R., & Uchiyama, K. (1999, April). *Professional development: A key to Kentucky's reform effort.* Paper presented at the annual meeting of the American Educational Research Association. Montreal, PQ, Canada.

Cohen, D. (1990, Fall). A revolution in one classroom: The case of Mrs. Oublier. *Educational Evaluation and Policy Analysis, 12,* 327-345.

Ensign, G. (1998, May). *The Washington Assessment of Student Learning, an update, May 1998.* Seattle, WA: Commission on Student Learning. Retrieved February 28, 2000, from http://www.k12.wa.us/reform/assess/subdocuments/asesorerv.html

Goertz, M., Duffy, M., & Le Floch, K. (2001, March). *Assessment and accountability systems in 50 states: 1999-2000* (RR-046). Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education.

Herman, J. L., Klein, D. C. D., & Abedi, J. (2000). Assessing students' opportunity to learn: Teacher and student perspectives. *Educational Measurement: Issues and Practice, 19*(4), 16-24

Jerald, C. D. (2000, January). The state of the states. *Quality counts 2000. Who should teach?* Bethesda, MD: Education Week. Retrieved 19 October 2001 from http://www.edweek.org/sreports/qc00/templates/article.cfm?slug=sosintro.htm

Koretz, D. M., Barron, S., Mitchell, K. J., & Stecher, B. M. (1996). *Perceived effects of the Kentucky Instructional Results Information System (KIRIS)* (MR-792-PCT/FF). Santa Monica, CA: RAND.

McIver, M. C., & Wolf, S. A. (1998, November). *Writing conferences: Powerful tools for writing instruction* (CSE Tech. Rep. No. 494). Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing (CRESST).

State of Washington Joint Legislative and Audit Review Committee. (1999). K-12 Finance and student performance study (Report 99-9, September 15). Olympia, WA: Author.

Stecher, B. (1999, September). *The impact of Kentucky's accountability system on instruction and outcomes.* Presentation at the annual CRESST conference, Benchmarks for Accountability: Are We There Yet?, Los Angeles, University of California.

Stecher, B., & Barron, S. (1999). *Quadrennial milepost accountability testing in Kentucky* (CSE Tech. Rep. No. 505). Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Stecher, B., Barron, S. E., Chun, T., & Ross, K. (2000, August). *The effects of Washington education reform on schools and classrooms* (CSE Tech. Rep. No. 525). Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Stecher, B., Barron, S. E., Kaganoff, T., & Goodwin, J. (1998, June). *The effects of standards-based assessment on classroom practices: Results of the 1996-97 RAND survey of Kentucky teachers of mathematics and writing* (CSE Tech. Rep. No. 482). Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Stecher, B., & Mitchell, K. (1995) *Portfolio-driven reform: Vermont teachers' understanding of mathematical problem solving* (CSE Tech. Rep. No. 400) Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Student Learning and Improvement Act, Rev. Code of Wash. § 28A.150.210 (1993).

Wolf, S. A., & McIver, M. C. (1999). When progress becomes policy: The paradox of Kentucky state reform for exemplary teachers. *Phi Delta Kappan, 80,* 401-406.

Appendix

Table A.1

Definition of Mathematics Practice Composite Variables

Composite variable	Items
Mathematical problem solving (8 items) alpha = .92 (4th), .92 (7th)	Frequency: formulate questions, construct solutions, analyze information, draw conclusions and verify results, gather information, organize and interpret information, represent and share information, and predict results and make inferences.
Relating mathematics to life and other fields (5 items) alpha = .85 (4th), .80 (7th)	Frequency: relate mathematical concepts and procedures to other disciplines, relate mathematical concepts to real-life situations, give examples of real-life applications of mathematics skills, use mathematics to solve real-life problems, use mathematics in the context of other subjects.
Constructive mathematics (4 items) alpha = .80 (4th), .81 (7th)	Frequency: demonstrate mathematical ideas using constructions, manipulatives, etc., solve problems using manipulatives, work on extended mathematics activities that take several days, discover mathematics concepts for themselves.
Focus on mathematics EALRs and rubrics (4 items) alpha = .78 (4th) .81 (7th)	Frequency: discuss EALRs in mathematics with your students, have students score classroom work using mathematics rubrics, use open-ended questions in classroom work, display scoring rubrics in classroom.
Focus on measurement (4 items) alpha = .74 (4th) .75 (7th)	Frequency: measurement, geometric sense, probability and statistics, use measuring tools in mathematics.
Practice WASL in mathematics (2 items) alpha = .76 (4th), .90 (7th)	Frequency: have students practice using WASL-like items, discuss responses to WASL or WASL-like items that illustrate different levels of performance.
Demonstrating mathematics (2 items) alpha = .76 (4th), .86 (7th)	Frequency: demonstrate how to perform a new mathematics skill, explain a new concept.
Peer interaction in mathematics (2 items) alpha = .78 (4th), .78 (7th)	Frequency: [students] explain their thinking to other students, work on problem solving in groups with other students.

Table A.2

Definition of Writing Practice Composite Variables

Composite variable	Items
Writing process (5 items) alpha = .91 (4th), .94 (7th)	Frequency: pre-write, draft, revise, edit, publish.
Conference and assessment of writing (6 items) alpha = .84 (4th) .83 (7th)	Frequency: assessment of student's own strength and needs for improvement, seek and offer feedback to other students, give example of choosing appropriate words to describe objects or experiences, conference with students about their writing, provide time for students to conference with each other about writing, assess students' writing skills.
Focus on theme, audience and purpose (5 items) alpha = .88 (4th), .87 (7th)	Frequency: development of concept and design, style appropriate to audience and purpose, write for different audiences, write for different purposes, write in a variety of forms.
Rubric-based writing (4 items) alpha = .83 (4th), .74 (7th)	Frequency: teach Six-Trait or other rubric-based approach to writing, have students score classroom work using rubrics, use open-ended questions in classroom work, display scoring rubrics in classroom.
Practice WASL in writing (4 items) alpha = .78 (4th), .76 (7th)	Frequency: comment on student writing in terms of WASL rubrics, provide a prompt to initiate student writing, have students practice using WASL-like items, discuss responses to WASL or WASL-like items that illustrate different levels of performance.
Focus on writing conventions (2 items) alpha =. 65 (4th), .74 (7th)	Frequency: application of writing conventions, explain correct usage of grammar, spelling, punctuation, and syntax.

Table A.3

Definition of Teacher Knowledge and Opinion Composite Variables

Composite variable	Items
Professional development about reform (3 items) alpha = .72 (4th), .83 (7th)	Amount: professional development related to EALRs, WASL, and alignment of curriculum and instruction with EALRs.
Understanding of reform (3 items) alpha = .75 (4th) .82 (7th)	Level of understanding: about EALRs, WASL, and alignment of curriculum and instruction with EALRs.
WASL is appropriate and useful (6 items) alpha = .78 (4th), .73 (7th)	Agreement: WASL standards for math are appropriate, WASL standards for writing are appropriate, it is easy to raise WASL scores, WASL classroom score reports are useful, WASL school information is useful, and information from state testing is useful.

Table A.4

Definition of Principal Knowledge and Opinion Composite Variables

Composite variable	Items
Understanding of reform (3 items) alpha = .80 (4th) ,86 (7th)	Level of understanding: about EALRs, WASL, and alignment of curriculum and instruction with EALRs.
District standards aligned with EALRs (4 times) alpha = .90 (4th), .86 (7th)	Degree of alignment: reading, writing, mathematics, communication.
School curriculum aligned with EALRs (4 items) alpha = .79 (4th), .76 (7th)	Degree of alignment: reading, writing, mathematics, communication.
School still needs to work on reform (5 items) alpha = .62 (4th), .74 (7th)	Agreement: teachers find it difficult to align curriculum with EALRs, teachers are slow to change practices to support reform, teachers do not understand content they need to know, teachers do not all feel responsible for improving WASL scores, WASL strained my relationship with teachers.
Support for reform (4 items) alpha = .62 (4th), .71 (7th)	Agreement: goals of reform are attainable, WASL standards in mathematics are appropriate, WASL standards in writing are appropriate, do not expect new reform to replace WASL/EALRs in five years.
Tests provide useful information (3 items) alpha = .76 (4th), .65 (7th)	Agreement: WASL classroom score reports are useful, WASL school reports are useful, the information learned from tests is worth the cost.
School actions to support reform	Number of policies enacted to support reform (of 21 possible).

Table A.5

Mean Teacher Frequency Scores on Mathematics Practice Composite Variables by School Percent Free or Reduced-Price Lunch

Composite variable	Fourth grade		Seventh grade	
	Low[a]	High[b]	Low[c]	High[d]
Focus on number sense	4.3*	4.0*	3.9	4.1
Direct instruction	4.2	4.0	4.0**	4.5**
Peer interaction	4.0	3.8	3.5	3.5
Problem solving	3.7	3.7	3.4	3.7
Write about mathematics	3.6	3.7	3.5	3.6
Relating mathematics to other areas	3.5	3.3	3.6	3.5
Practice WASL	3.4	3.6	3.1	3.5
Conduct speed drills	3.4	3.5	2.0	1.9
Constructive mathematics	3.2	3.0	2.7	2.9
Focus on measurement	3.0	2.8	3.0	2.9
Focus on EALRs and rubrics	2.8	2.8	2.4*	2.9*

[a]Less than or equal to 33.8% free/reduced-price lunch. [b]Greater than 33.8% free/reduced-price lunch. [c]Less than or equal to 29.1% free/reduced-price lunch. [d]Greater than 29.1% free/reduced-price lunch.

*$p < .05$. ** $p < .01$.

Table A.6

Mean Teacher Frequency Scores on Writing Practice Composite Variables by School Percent Free or Reduced-Price Lunch

Composite variable	Fourth grade		Seventh grade	
	Low[a]	High[b]	Low[c]	High[d]
Writing process	3.7	3.6	3.9	3.6
Conference and assessment	3.4*	3.1*	3.3	3.2
Focus on theme, audience and purpose	3.5**	3.2**	3.7	3.4
Rubric-based writing	3.4*	3.1*	3.4	3.3
Practice WASL	3.1	3.1	3.1	3.2
Focus on conventions	4.3	4.0	4.3	4.4

[a]Less than or equal to 33.8% free/reduced-price lunch. [b]Greater than 33.8% free/reduced-price lunch. [c]Less than or equal to 29.1% free/reduced-price lunch. [d]Greater than 29.1% free/reduced-price lunch.

*$p < .05$. ** $p < .01$.

Table A.7

Regression of Fourth- and Seventh-Grade WASL Mathematics Scores on School Demographic and Composite Principal and Teacher Variables

Composite variables	Fourth grade		Seventh grade	
	Beta	PR > *F*	Beta	PR > *F*
Intercept	0	0.748	0	0.498
Percent free/reduced-price lunch	-0.308	0.008	-0.702	<.0001
Percent Asian	0.227	0.026	0.326	0.002
Percent American Indian	-0.175	0.082		
Percent Black	-0.173	0.087	-0.191	0.080
Percent Hispanic	-0.409	0.001		
Percent female				
Enrollment				
Principal composite variables				
Understanding of reform				
School curriculum aligned with EALRs			0.187	0.043
School still needs to work on reform	-0.272	0.005		
Principal support for reform				
Tests provide useful information				
School actions to support reform				
Teacher composite variables				
Professional development about reform				
Understanding of reform				
WASL is appropriate and useful				
Mathematics practice composite variables				
Mathematical problem solving				
Relating mathematics to life and other fields	0.324	0.003		
Constructive mathematics				
Focus on mathematics EALRs and rubrics				
Focus on measurement	-0.164	0.120		
Practice WASL in mathematics				
Demonstrating mathematics	0.162	0.086		
Peer Interaction in mathematics				
R-squared	.70		.76	

Note. Final stepwise regression model (0.15 significance level for entry).

Table A.8

Regression of Fourth- and Seventh-Grade WASL Writing Scores on School Demographic and Composite Principal and Teacher Variables

Composite variables	Fourth grade		Seventh grade	
	Beta	PR > *F*	Beta	PR > *F*
Intercept	0	0.235	0	0.131
Percent free/reduced-price lunch	-0.292	0.015	-0.604	<.0001
Percent Asian	0.249	0.034		
Percent American Indian				
Percent Black				
Percent Hispanic				
Percent female				
Enrollment				
Principal composite variables				
Understanding of reform			0.221	0.060
School curriculum aligned with EALRs				
School still needs to work on reform	-0.179	0.115		
Principal support for reform	-0.234	0.040		
Tests provide useful information				
School actions to support reform			-0.324	0.004
Teacher composite variables				
Professional development about reform				
Understanding of reform				
WASL is appropriate and useful			0.160	0.141
Writing practice composite variables				
Writing process				
Conference and assessment of writing				
Focus on theme, audience and purpose				
Rubric-based writing	0.270	0.024		
Practice WASL in writing				
Focus on writing conventions			0.285	0.014
R-squared	.45		.66	

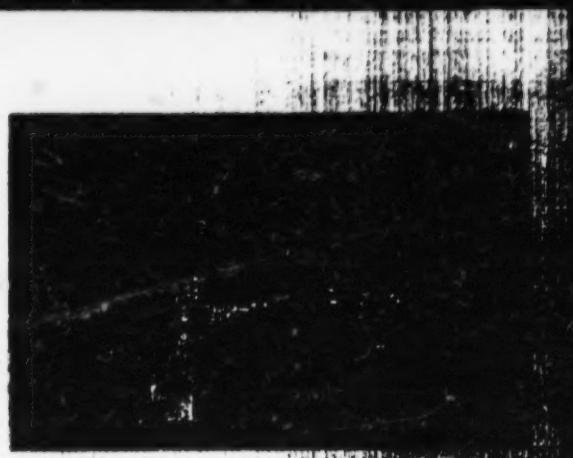
Note. Final stepwise regression model (0.15 significance level for entry).

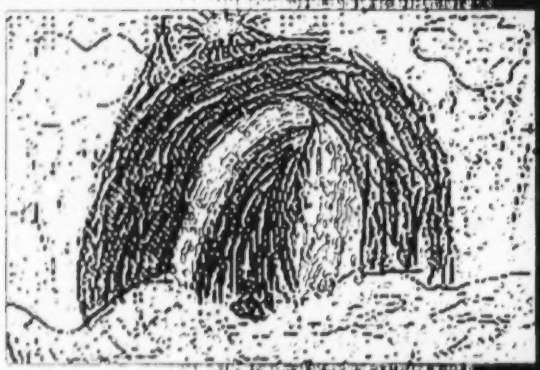
Table A.9

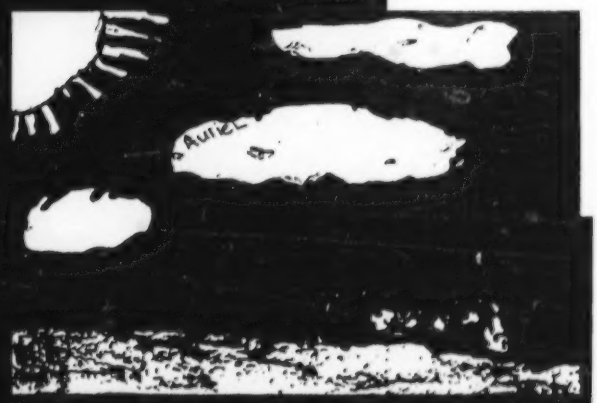
Regression of Fourth- and Seventh-Grade WASL Reading Scores on School Demographic and Composite Principal and Teacher Variables

Composite variables	Fourth grade		Seventh grade	
	Beta	PR > *F*	Beta	PR > *F*
Intercept	0	0.046	0	0.778
Percent free/reduced-price lunch	-0.312	0.010	-0.345	0.006
Percent Asian				
Percent American Indian	-0.382	0.015		
Percent Black			-0.162	0.089
Percent Hispanic	-0.382	0.002	-0.562	<.0001
Percent female				
Enrollment	-0.172	0.092		
Principal composite variables				
Understanding of reform				
School curriculum aligned with EALRs			0.160	0.094
School still needs to work on reform	-0.185	0.058		
Principal support for reform				
Tests provide useful information				
School actions to support reform				
Teacher composite variables				
Professional development about reform	0.155	0.133		
Understanding of reform				
WASL is appropriate and useful				
R-squared	.61		.73	

Note. Final stepwise regression model (0.15 significance level for entry).







BEST COPY AVAILABLE